

# REBASE – restriction enzymes and methylases

Richard J. Roberts\* and Dana Macelis

New England BioLabs, 32 Tozer Road, Beverly, MA 01915, USA

Received November 6, 1997; Accepted November 6, 1997

## ABSTRACT

**REBASE is a comprehensive database of information about restriction enzymes and their associated methylases, including their recognition and cleavage sites and their commercial availability. Also included is a listing of homing endonucleases. Information from REBASE is available via monthly electronic mailings as well as via anonymous ftp and through the World Wide Web. The REBASE web site, <http://www.neb.com/rebase>, is where we maintain a web page for every enzyme, reference and supplier. Additionally, there is a search facility, help and NEWS pages, and a complete description of our various services. Specialized files are available that can be used directly by many software packages.**

## INTRODUCTION

The restriction enzyme database, REBASE, is a collection of information about restriction enzymes and methylases. Since the last description of the contents of REBASE (1), 79 new entries have been added. These include two Type II enzymes with new specificities: *Bse*MII which recognizes the sequence CTSAG (2), and *Hae*IV, which recognizes the sequence GAYNNNNNRTC (3). Among the 2900 restriction enzymes now known, there are 19 fully characterized Type I specificities, 212 different Type II specificities and four different Type III specificities. Table 1 contains a listing of all prototype restriction enzymes (Types I, II and III), together with their commercially available isoschizomers and neoschizomers. Homing endonucleases (4) are also listed in REBASE and seven of the 37 known are commercially available (see Table 2). REBASE currently stores 4696 published references, with abstracts, and 486 unpublished observations.

REBASE is updated daily. Each month, a set of REBASE data files are released publically and distributed to the scientific community, at no charge, via e-mail. They can also be retrieved by anonymous ftp ([www.neb.com](http://www.neb.com), in `pub/rebase`), WAIS ([www.neb.com](http://www.neb.com), port 210) or through the World Wide Web (<http://www.neb.com/rebase>). These data files are ASCII text files, many of which are designed specifically for use with a variety of software packages such as GCG, IGSuite, GENEPRO, Staden, DNA Strider, Pro-Cite, PC/Gene, SEQAIDII, GENE-TYX, DNASIS, CAD Gene and Reference Manager. Other data files include a complete set of references, including abstracts, to papers on restriction enzymes and methylases; a list of all

commercial suppliers of restriction enzymes and methylases, complete with contact information and a listing of enzymes they currently sell. New data files are constantly being added and each release of REBASE includes a monthly newsletter indicating that the files at the `www` and `ftp` sites have been updated, and listing new enzymes, newly available formats, enzyme name changes, etc. To join the mailing list or for more information, contact either author (e-mail: [macelis@neb.com](mailto:macelis@neb.com) or [roberts@neb.com](mailto:roberts@neb.com); telephone (978) 927–3382; fax (978) 921–1527).

Because of the new naming conventions for homing endonucleases (4) several enzyme names have been changed. Those affected are the homing endonucleases NOT encoded in introns. These now carry the prefix F- and are called ‘free-standing endonucleases’.

Old name	New name
Endo. <i>Sce</i> I	F- <i>Sce</i> I
HO-endo	F- <i>Sce</i> II
I- <i>Suv</i> I	F- <i>Suv</i> I
N- <i>Tev</i> I	F- <i>Tev</i> I
N- <i>Tev</i> II	F- <i>Tev</i> II

In forming Table 1, all endonucleases cleaving DNA at a specific sequence have been considered to be restriction enzymes, although in most cases there is no direct genetic evidence for the presence of a restriction-modification system. The endonucleases are named in accordance with the proposal of Smith and Nathans (5).

We request that authors wishing to cite REBASE use this article as a general reference for the complete listing that is available electronically.

## ACKNOWLEDGEMENTS

Special thanks are due to the many individuals who have so kindly communicated their unpublished results for inclusion in this compilation. This database is supported by the National Library of Medicine (LM04971).

## REFERENCES

- 1 Roberts, R.J. and Macelis, D. (1997) *Nucleic Acids Res.*, **25**, 248–262.
- 2 Kazlauskienė, R., Petrusyte, M., Kairaitienė, R., Kiuduliene, E., Butkus, V. and Janulaitis, A. unpublished observations.
- 3 Piekarowicz, A. unpublished observations.
- 4 Belfort, M. and Roberts, R.J. (1997) *Nucleic Acids Res.*, **25**, 3379–3388.
- 5 Smith, H.O. and Nathans, D. (1973) *J. Mol. Biol.*, **81**, 419–423.

\* To whom correspondence should be addressed. Tel: +1 978 927 3382; Fax: +1 978 921 1527; Email: [roberts@neb.com](mailto:roberts@neb.com)

**Table 1**  
**Type II restriction enzymes**

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
<i>Aat</i> II		GACGT↓C		ADEFKLMNOPRS
<i>Ace</i> I		GT↓MKAC	5(6)	ABDEGJKLMNQPORS
( <i>Ace</i> III)		CAGCTC(7/11)		
<i>Ac</i> I		CCGC(-3/-1)		N
<i>Ac</i> II		AA↓CGTT		IN
	<i>Psp</i> 1406I	AA↓CGTT		DFKM
<i>Acy</i> I		GR↓CGYC		JM
	<i>Bbi</i> II	GR↓CGYC		K
	<i>Bsa</i> HI	GR↓CGYC		N
	<i>Bst</i> ACI	GR↓CGYC		I
	<i>Hin</i> 1I	GR↓CGYC		EFO
	<i>Hsp</i> 92I	GR↓CGYC		R
	<i>Msp</i> 17I	GR↓CGYC		D
<i>Af</i> III		C↓TTAAG		ABJKNO
	<i>Bfr</i> I	C↓TTAAG		MO
	<i>Bsp</i> TI	C↓TTAAG		DF
	<i>Bst</i> 98I	C↓TTAAG		R
	<i>Msp</i> CI	C↓TTAAG		C
	<i>Vha</i> 464I	C↓TTAAG		I
<i>Af</i> III		A↓CRYGT		ABMN
<i>Age</i> I		A↓CCGGT	?(5)	JNR
	<i>Asi</i> AI	A↓CCGGT		I
	<i>Pin</i> AI	A↓CCGGT		BM
( <i>Aha</i> III)	<i>Dra</i> I	TTT↓AAA		ABDEFGHIJKLMNQPORS
<i>Alu</i> I		AG↓CT	3(5)	ABCDEFGHIJKLMNQPQRST
<i>Alw</i> NI		CAGNNN↓CTG		ABN
<i>Apa</i> I		GGGCC↓C	4(5)	ABEGIKLMNQPORS
	<i>Bsp</i> 120I	G↓GGCCC		DFN
	<i>Psp</i> OMI	G↓GGCCC		I
( <i>Apa</i> BI)	<i>Bst</i> API	GCANNNN↓NTGC		IN
<i>Apa</i> LI		G↓TGCAC	?(5)	ADEKN
	<i>Alw</i> 44I	G↓TGCAC		FJMORS
	<i>Vne</i> I	G↓TGCAC		I
<i>Ap</i> oI		R↓AATTY	?(6)	N
	<i>Acs</i> I	R↓AATTY		DIM
<i>Asc</i> I		GG↓CGCGCC	?(5)	N
( <i>Asu</i> I)	<i>Asp</i> S9I	G↓GNCC		I
	<i>Bsi</i> ZI	G↓GNCC		T
	<i>Cfr</i> 13I	G↓GNCC	4(5)	DFKO
	<i>Sau</i> 96I	G↓GNCC	4(5)	AJLMNOR
	( <i>Fmu</i> I)	GGNC↓C		
( <i>Asu</i> II)	<i>Bpu</i> 14I	TT↓CGAA		I
	<i>Bsi</i> CI	TT↓CGAA		T
	<i>Bsp</i> 119I	TT↓CGAA		DF
	<i>Bst</i> BI	TT↓CGAA		N
	<i>Psy</i> I	GACN↓NNGTC		
	<i>Cbi</i> I	TT↓CGAA		J
	<i>Csp</i> 45I	TT↓CGAA		OR
	<i>Lsp</i> I	TT↓CGAA		L

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>NspV</i>	TT↓CGAA		ABJKO
	<i>SfiI</i>	TT↓CGAA		M
<i>AvaI</i>		C↓YCGRG	?(4)	ABEGJKLMNQPORS
	<i>Ama87I</i>	C↓YCGRG		DI
	<i>BcoI</i>	C↓YCGRG		AT
	<i>BsoBI</i>	C↓YCGRG	?(4)	N
	<i>Eco88I</i>	C↓YCGRG		F
	( <i>Nli387/7I</i> )	CYCGR↓G		
<i>AvaII</i>		G↓GWCC	?(5)	ABEGJKMNPQRS
	<i>Bme18I</i>	G↓GWCC		I
	<i>Eco47I</i>	G↓GWCC		FO
	<i>HgiEI</i>	G↓GWCC	?(5)	D
	<i>SinI</i>	G↓GWCC	4(5)	LR
	( <i>VpaK11AI</i> )	↓GGWCC		
( <i>AvaIII</i> )	<i>EcoT22I</i>	ATGCA↓T		AKO
	<i>Mph1103I</i>	ATGCA↓T		F
	<i>NsiI</i>	ATGCA↓T		BDEHJLMNRS
	<i>Zsp2I</i>	ATGCA↓T		I
	<i>Ppu10I</i>	A↓TGCAT		DN
	( <i>BfrBI</i> )	ATG↓CAT		
<i>AvrII</i>		C↓CTAGG		N
	<i>BlnI</i>	C↓CTAGG		AKMS
	<i>BspA2I</i>	C↓CTAGG		I
( <i>BaeI</i> ) <sup>e</sup>		(10/15)ACNNNNGTAYC(12/7)	1(6),-4(6)	
<i>BalI</i>		TGG↓CCA	4(4)	AJKR
	<i>MlsI</i>	TGG↓CCA		F
	<i>MluNI</i>	TGG↓CCA		MS
	<i>MscI</i>	TGG↓CCA		BDNOS
<i>BamHI</i>		G↓GATCC	5(4)	ABCDEFGHIJKLMNQPQRST
<i>BbvI</i>		GCAGC(8/12)	2(5),-2(5)	N
	<i>Bst7II</i>	GCAGC(8/12)		R
( <i>BbvII</i> )	<i>BbsI</i>	GAAGAC(2/6)		N
	<i>BpiI</i>	GAAGAC(2/6)		DF
	<i>BpuAI</i>	GAAGAC(2/6)		M
( <i>BccI</i> )		CCATC		
( <i>Bce83I</i> )		CTTGAG(16/14)		
( <i>BceI</i> )		ACGGC(12/13)		
<i>BcgI</i> <sup>5</sup>		(10/12)GCANNNNNNTCG(12/10)	3(6),-3(6)	N
<i>BciVI</i>		GTATCC(6/5)		N
<i>BclI</i>		T↓GATCA		BCDEFGJLMNOPRS
	<i>BsiQI</i>	T↓GATCA		T
	<i>FbaI</i>	T↓GATCA		AK
	<i>Ksp22I</i>	T↓GATCA		I
( <i>BetI</i> )	<i>BsaWI</i>	W↓CCGGW		N
( <i>BfiI</i> )		ACTGGG(4/5)		
<i>BglI</i>		GCCNNNN↓NGGC	?(4)	ABCDEFGHIJKLMNQPQRST
<i>BglII</i>		A↓GATCT	5(4)	ABCDEFGHIJKLMNQPQRST
( <i>BinI</i> )	<i>ActWI</i>	GGATC(4/5)		I
	<i>AlwI</i>	GGATC(4/5)		N
( <i>BmgI</i> )		GKGCCC		

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
( <i>Bpl</i> I)		GAGNNNNNCTC		
<i>Bpu</i> 10I		CCTNAGC(-5/-2)		F
<i>Bsa</i> AI		YAC↓GTR		N
	<i>Bst</i> BAI	YAC↓GTR		I
<i>Bsa</i> BI		GATNN↓NNATC		N
	<i>Bse</i> 8I	GATNN↓NNATC		I
	<i>Bsh</i> 1365I	GATNN↓NNATC		DF
	<i>Bsi</i> BI	GATNN↓NNATC		T
	<i>Bsr</i> BRI	GATNN↓NNATC		R
	<i>Mam</i> I	GATNN↓NNATC	?(6)	M
( <i>Bsa</i> XI)		ACNNNNNCTCC		
( <i>Bsb</i> I)		CAACAC		
( <i>Bsc</i> GI)		CCCGT		
( <i>Bse</i> MII)		CTSAG		
<i>Bse</i> PI		G↓CGCGC		I
	<i>Bss</i> HII	G↓CGCGC	?(5)	ABDEJKLMNOPS
	<i>Paul</i> I	G↓CGCGC		F
<i>Bse</i> RI		GAGGAG(10/8)		N
<i>Bsg</i> I		GTGCAG(16/14)		N
( <i>Bsi</i> I)	<i>Bss</i> SI	CACGAG(-5/-1)		N
	<i>Bst</i> 2BI	CACGAG(-5/-1)		DI
<i>Bsi</i> YI		CCNNNNN↓NNGG		MT
	<i>Bsc</i> 4I	CCNNNNN↓NNGG		I
	<i>Bse</i> LI	CCNNNNN↓NNGG		F
	<i>Bsl</i> I	CCNNNNN↓NNGG		N
<i>Bsm</i> I		GAATGC(1/-1)		ABDEJLMNOS
	<i>Bsa</i> MI	GAATGC(1/-1)		R
	<i>Bsc</i> CI	GAATGC(1/-1)		T
	<i>Mva</i> 1269I	GAATGC(1/-1)		F
<i>Bsm</i> AI		GTCTC(1/5)		N
	<i>Alw</i> 26I	GTCTC(1/5)	3(5),-4(6)	DFR
( <i>Bsp</i> 24I) <sup>e</sup>		(8/13)GACNNNNNNTGG(12/7)		
<i>Bsp</i> 1407I		T↓GTACA		DFK
	<i>Aau</i> I	T↓GTACA		I
	<i>Bsr</i> GI	T↓GTACA		N
	<i>Ssp</i> BI	T↓GTACA		M
( <i>Bsp</i> GI)		CTGGAC		
<i>Bsp</i> HI		T↓CATGA		AN
	<i>Pag</i> I	T↓CATGA		F
	<i>Rca</i> I	T↓CATGA		BM
<i>Bsp</i> LU11I		A↓CATGT		M
<i>Bsp</i> MI		ACCTGC(4/8)		N
( <i>Bsp</i> MII)	<i>Acc</i> III	T↓CCGGA		EJKQR
	<i>Bse</i> AI	T↓CCGGA		CM
	<i>Bsi</i> MI	T↓CCGGA		AT
	<i>Bsp</i> 13I	T↓CCGGA		I
	<i>Bsp</i> EI	T↓CCGGA		N
	<i>Kpn</i> 2I	T↓CCGGA		BDF
	<i>Mro</i> I	T↓CCGGA		MO
<i>Bsr</i> I		ACTGG(1/-1)		N

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>Bse</i> I	ACTGG(1/-1)		I
	<i>Bse</i> NI	ACTGG(1/-1)		DF
	<i>Bsr</i> SI	ACTGG(1/-1)		R
<i>Bsr</i> BI		CCGCTC(-3/-3)		N
	<i>Acc</i> BSI	CCGCTC(-3/-3)		DI
	<i>Mb</i> I	CCGCTC(-3/-3)		F
<i>Bsr</i> DI		GCAATG(2/0)		N
	<i>Bse</i> 3DI	GCAATG(2/0)		I
<i>Bst</i> EII		G↓GTNACC		BDGHJLMNORS
	<i>Bst</i> PI	G↓GTNACC		K
	<i>Eco</i> 91I	G↓GTNACC		F
	<i>Eco</i> O65I	G↓GTNACC		AK
	<i>Psp</i> EI	G↓GTNACC		I
<i>Bst</i> XI		CCANNNNN↓NTGG		ABDEFGHJKLMNOQRS
<i>Cac</i> 8I		GCN↓NGC		N
( <i>Cau</i> II)	<i>Asu</i> C2I	CC↓SGG		I
	<i>Bcn</i> I	CC↓SGG	2(4)	ADFK
	<i>Nci</i> I	CC↓SGG		BEJLMNORST
	( <i>Eco</i> HI)	↓CCSGG		
<i>Cfr</i> I		Y↓GGCCR	4(5)	F
	<i>Eae</i> I	Y↓GGCCR	4(5)	ADKLMN
<i>Cfr</i> 10I		R↓CCGGY	2(5)	ADFKMO
	<i>Bse</i> 118I	R↓CCGGY		I
	<i>Bsr</i> FI	R↓CCGGY		N
	<i>Bss</i> AI	R↓CCGGY		C
( <i>Cje</i> I) <sup>5</sup>		(8/14)CCANNNNNNGT(15/9)		
( <i>Cje</i> PI) <sup>5</sup>		(7/13)CCANNNNNNNTC(14/8)		
<i>Cla</i> I		AT↓CGAT	5(6)	ABHKMNPRST
	<i>Ban</i> III	AT↓CGAT		O
	<i>Bsa</i> 29I	AT↓CGAT		I
	<i>Bsc</i> I	AT↓CGAT		L
	<i>Bse</i> CI	AT↓CGAT	5(6)	C
	<i>Bsi</i> XI	AT↓CGAT		T
	<i>Bsp</i> 106I	AT↓CGAT	5(6)	E
	<i>Bsp</i> DI	AT↓CGAT		N
	<i>Bsp</i> XI	AT↓CGAT		G
	<i>Bsu</i> 15I	AT↓CGAT	5(6)	DF
<i>Cvi</i> II		RG↓CY	3(5)	Q
( <i>Cvi</i> RI)		TG↓CA	4(6)	
<i>Dde</i> I		C↓TNAG	1(5)	ABEGLMNORS
	<i>Bst</i> DEI	C↓TNAG		DI
<i>Dpn</i> I*		G <sup>m6</sup> A↓TC		ABELMNRS
<i>Dra</i> II		RG↓GNCCY		AGMS
	<i>Eco</i> O109I	RG↓GNCCY		ABDEFJKLN
	( <i>Pss</i> I)	RGGNC↓CY		
<i>Dra</i> III		CACNNN↓GTG		AEGMNS
	<i>Ade</i> I	CACNNN↓GTG		F
<i>Drd</i> I		GACNNNN↓NNGTC		N
	<i>Dse</i> DI	GACNNNN↓NNGTC		I
( <i>Drd</i> II)		GAACCA		

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
<i>DsaI</i>		C↓CRYGG		M
	<i>BstDSI</i>	C↓CRYGG		DI
<i>Eam1105I</i>		GACNNN↓NNGTC		DFK
	<i>AhdI</i>	GACNNN↓NNGTC		N
	<i>AspEI</i>	GACNNN↓NNGTC		M
	<i>EclHki</i>	GACNNN↓NNGTC		R
	<i>NruGI</i>	GACNNN↓NNGTC		I
( <i>EciI</i> )		TCCGCC		
<i>Eco31I</i>		GGTCTC(1/5)	4(5),-4(6)	DF
	<i>BsaI</i>	GGTCTC(1/5)		N
<i>Eco47III</i>		AGC↓GCT		BDEFLMNORS
	<i>AfeI</i>	AGC↓GCT		I
	<i>Aor51HI</i>	AGC↓GCT		AK
<i>Eco57I</i>		CTGAAG(16/14)	5(6),-5(6)	DFN
<i>EcoNI</i>		CCTNN↓NNNAGG		AN
	<i>XagI</i>	CCTNN↓NNNAGG		F
<i>EcoRI</i>		G↓AATTC	3(6)	ABCDEFGHIJKLMNOPQRST
<i>EcoRII<sup>f</sup></i>		↓CCWGG	2(5)	BEJMOS
+	<i>BstLI</i>	CC↓WGG		T
+	<i>Bst2UI</i>	CC↓WGG		I
+	<i>BstNI</i>	CC↓WGG	2(4)	CEN
+	<i>BstOI</i>	CC↓WGG		R
+	<i>MvaI</i>	CC↓WGG	2(4)	ADFKMOS
<i>EcoRV</i>		GAT↓ATC	2(6)	ABCDEFGHIJKLMNOPQRST
	<i>Eco32I</i>	GAT↓ATC		F
( <i>EspI</i> )	<i>BlpI</i>	GC↓TNAGC		N
	<i>Bpu1102I</i>	GC↓TNAGC		BDEFK
	<i>Bsp1720I</i>	GC↓TNAGC		I
	<i>CelII</i>	GC↓TNAGC		ALM
<i>Esp3I</i>		CGTCTC(1/5)	4(5),-4(6)	DF
	<i>BsmBI</i>	CGTCTC(1/5)		N
<i>FauI</i>		CCCGC(4/6)		I
<i>FinI</i>		GGGAC		
	<i>BsmFI</i>	GGGAC(10/14)		N
<i>Fnu4HI</i>		GC↓NGC		N
	<i>Fsp4HI</i>	GC↓NGC		I
	<i>ItaI</i>	GC↓NGC		M
( <i>FnuDII</i> )	<i>AccII</i>	CG↓CG		AJKQ
	<i>Bsh1236I</i>	CG↓CG		DEF
	<i>BstUI</i>	CG↓CG		N
	<i>MvnI</i>	CG↓CG		M
	<i>ThaI</i>	CG↓CG		B
	( <i>SeI</i> )	↓CGCG		
<i>FokI</i>		GGATG(9/13)	3(6),-2(6)	ADGIJLMNR
	<i>BseGI</i>	GGATG(2/0)		F
	<i>BstF5I</i>	GGATG(2/0)	3(6)	IN
	( <i>StsI</i> )	GGATG(10/14)	3(6),-2(6)	
<i>FseI</i>		GGCCGG↓CC	?(5)	AKN
( <i>GdiIII</i> )		CGGCCR(-5/-1)		
<i>GsuI</i>		CTGGAG(16/14)		DF

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>BpmI</i>	CTGGAG(16/14)		N
( <i>HaeI</i> )		WGG↓CCW		
<i>HaeII</i>		RGCGC↓Y		ABDEGJKLMNORS
	<i>Bsp143II</i>	RGCGC↓Y		F
	<i>BstH2I</i>	RGCGC↓Y		I
	( <i>LpnI</i> )	RGC↓GCY		
<i>HaeIII</i>		GG↓CC	3(5)	ABCDGHIJKLMNOPQRST
	<i>BshI</i>	GG↓CC		T
	<i>BsuRI</i>	GG↓CC	3(5)	FI
	<i>PalI</i>	GG↓CC		E
( <i>HaeIV</i> )		GAYNNNNNRTC	2(6)	
<i>HgaI</i>		GACGC(5/10)	3(5),-2(5)	AN
( <i>HgiAI</i> )	<i>Abw21I</i>	GWGCW↓C		DF
	<i>AspHI</i>	GWGCW↓C		M
	<i>Bbv12I</i>	GWGCW↓C		I
	<i>BstHKA1</i>	GWGCW↓C		N
( <i>HgiCI</i> )	<i>AccB1I</i>	G↓GYRCC		I
	<i>BanI</i>	G↓GYRCC	?(5)	AEMNORS
	<i>BshNI</i>	G↓GYRCC		F
	<i>Eco64I</i>	G↓GYRCC		DF
( <i>HgiEII</i> )		ACCNNNNNNGGT		
( <i>HgiIII</i> )	<i>BanII</i>	GRGCY↓C	4(5)	ABKLMNOQRS
	<i>Eco24I</i>	GRGCY↓C		DF
	<i>EcoT38I</i>	GRGCY↓C		J
	<i>FriOI</i>	GRGCY↓C		I
<i>HhaI</i>		GCG↓C	2(5)	ABGJKNOPR
	<i>AspLEI</i>	GCG↓C		I
	<i>CfoI</i>	GCG↓C		BLMRS
	<i>Hin6I</i>	G↓CGC		DF
	<i>HinP1I</i>	G↓CGC		N
	<i>HspAI</i>	G↓CGC		I
( <i>Hin4I</i> )		GABNNNNNVTC		
<i>HindII</i>		GTY↓RAC	5(6)	MS
	<i>HincII</i>	GTY↓RAC	5(6)	ABCDEFGHIJKLMNOPQR
<i>HindIII</i>		A↓AGCTT	1(6)	ABCDEFGHIJKLMNOPQRST
<i>HinII</i>		G↓ANTC		ABCDEFGHIJKLMNOPQRST
<i>HpaI</i>		GTT↓AAC	5(6)	ABDEGHJKLMNOPQRST
	<i>BstHPI</i>	GTT↓AAC		I
	<i>KspAI</i>	GTT↓AAC		F
<i>HpaII</i>		C↓CGG	2(5)	BDEFGILMNOPQRS
	<i>BsiSI</i>	C↓CGG		C
	<i>HapII</i>	C↓CGG	2(5)	AK
	<i>Hin2I</i>	C↓CGG		F
	<i>MspI</i>	C↓CGG	1(5)	ABCDEFGHIJKLMNOPQRS
<i>HphI</i>		GGTGA(8/7)	5(6),-2(5)	AFN
	<i>AsuHPI</i>	GGTGA(8/7)		I
<i>KpnI</i>		GGTAC↓C	4(6)	ABCDEFGHIJKLMNOPQRST
	<i>Acc65I</i>	G↓GTACC		DFINR
	<i>Asp718I</i>	G↓GTACC		M
<i>Ksp632I</i>		CTCTTC(1/4)		M

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>Bsu6I</i>	CTCTTC(1/4)		I
	<i>Eam1104I</i>	CTCTTC(1/4)		DEF
	<i>EarI</i>	CTCTTC(1/4)		N
<i>MaeI</i>		C↓TAG		M
	<i>BfaI</i>	C↓TAG		N
<i>MaeII</i>		A↓CGT		M
	<i>TaiI</i>	ACGT↓		FN
	<i>TscI</i>	ACGT↓		L
<i>MaeIII</i>		↓GTNAC		M
<i>MboI</i> <sup>g</sup>		↓GATC	2(6)	ABCEFGKLNQRT
	<i>BscFI</i>	↓GATC		T
+	<i>Bsp143I</i>	↓GATC		DF
	<i>DpnII</i>	↓GATC	2(6)	N
	<i>Kzo9I</i>	↓GATC		I
	<i>NdeII</i>	↓GATC		BDGJMS
+	<i>Sau3AI</i>	↓GATC	4(5)	ABEGHIJKLMNQPORS
+	<i>BspKT6I</i>	GAT↓C	2(6)	
	<i>ChaI</i>	GATC↓		
<i>MboII</i>		GAAGA(8/7)	5(6)	ABDFGJKNOQR
( <i>McrI</i> )	<i>BsaOI</i>	CGRY↓CG		R
	<i>Bsh1285I</i>	CGRY↓CG		DF
	<i>BsiEI</i>	CGRY↓CG		N
	<i>BstMCI</i>	CGRY↓CG		I
<i>MfeI</i>		C↓AATTG		N
	<i>MunI</i>	C↓AATTG	3(6)	BDEFKMS
<i>MluI</i>		A↓CGCGT		ABDEFGHIJKLMNQPORS
( <i>MmeI</i> )		TCCRAC(20/18)		
<i>MnlI</i>		CCTC(7/6)		AEFNQ
<i>MseI</i>		T↓TAA		BN
	<i>TruII</i>	T↓TAA		F
	<i>Tru9I</i>	T↓TAA		DILMR
<i>MsiI</i>		CAYNN↓NNRTG		N
( <i>MstI</i> )	<i>Acc16I</i>	TGC↓GCA		DI
	<i>AviII</i>	TGC↓GCA		M
	<i>FspI</i>	TGC↓GCA		ABJKNOS
	<i>NsbI</i>	TGC↓GCA		F
<i>MwoI</i>		GCNNNNN↓NNGC	?(5)	N
<i>NaeI</i>		GCC↓GGC		ADEKLMNORS
	<i>MroNI</i>	G↓CCGGC		I
	<i>NgøAIV</i>	G↓CCGGC		B
	<i>NgøMI</i>	G↓CCGGC	2(5)	NR
<i>NarI</i>		GG↓CGCC		BEJMNOPRS
	<i>BbeI</i>	GGCGC↓C		AK
	<i>EgeI</i>	GGC↓GCC		I
	<i>EheI</i>	GGC↓GCC		ADFNO
	<i>KasI</i>	G↓GCGCC		N
	<i>Mly113I</i>	GG↓CGCC		I
<i>NcoI</i>		C↓CATGG	?(4)	ABCDEFGHIJKLMNQPQRST
	<i>Bsp19I</i>	C↓CATGG		I
<i>NdeI</i>		CA↓TATG		ABDEFGJKLMNPRS

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
<i>NheI</i>	<i>Fau</i> NDI	CA↓TATG		DI
	<i>Asu</i> NHI ( <i>Ace</i> II)	G↓CTAGC		ABDEFGJKLMNOPRS
<i>Nla</i> III	<i>Hsp</i> 92II ( <i>Cvi</i> AII)	G↓CTAGC		I
		GCTAG↓C		
<i>Nla</i> IV	<i>Hsp</i> 92II ( <i>Cvi</i> AII)	CATG↓	2(6)	ANO
		CATG↓		R
<i>Nla</i> IV		C↓ATG	2(6)	
		GGN↓NCC		N
	<i>Bsc</i> BI	GGN↓NCC		T
	<i>Bsp</i> LI	GGN↓NCC		F
<i>Not</i> I	<i>Psp</i> N4I	GGN↓NCC		I
		GC↓GGCCGC	?(4)	ABCDEFGHIJKLMNOPQRST
<i>Nru</i> I	<i>Cci</i> NI	GC↓GGCCGC		I
		TCG↓CGA		ABCDEFGHIJKLMNOQRST
<i>Nsp</i> I	<i>Bsp</i> 68I	TCG↓CGA		F
		RCATG↓Y		ABKM
<i>Nsp</i> BII	<i>Bst</i> NSI	RCATG↓Y		DI
		CMG↓CKG		A
<i>Pac</i> I	<i>Msp</i> AII	CMG↓CKG		INR
		TTAAT↓TAA		NO
<i>Pfl</i> 1108I		TCGTAG		
<i>Pfl</i> MI		CCANNNN↓NTGG		ANS
	<i>Acc</i> B7I	CCANNNN↓NTGG		IR
	<i>Esp</i> 1396I	CCANNNN↓NTGG		F
	<i>Van</i> 91I	CCANNNN↓NTGG		DFKM
<i>Ple</i> I		GAGTC(4/5)		N
	( <i>Mly</i> I)	GACTC(5/5)		
<i>Pma</i> CI		CAC↓GTG		AK
	<i>Bbr</i> PI	CAC↓GTG		MO
	<i>Eco</i> 72I	CAC↓GTG		DEFR
	<i>Pml</i> I	CAC↓GTG		N
<i>Pme</i> I		GTTT↓AAAC		N
<i>Ppu</i> MI		RG↓GWCCY		ANO
	<i>Psp</i> 5II	RG↓GWCCY		BDF
	<i>Psp</i> PPI	RG↓GWCCY		I
<i>Psh</i> AI		GACNN↓NNGTC		AKN
<i>Pst</i> I		CTGCA↓G	5(6)	ABCDEFGHIJKLMNOPQRST
<i>Pvu</i> I		CGAT↓CG		ABDEFGJKLMNOPQRS
	<i>Bsp</i> CI	CGAT↓CG		E
	<i>Ple</i> 19I	CGAT↓CG		I
<i>Pvu</i> II ( <i>Rle</i> AI)		CAG↓CTG	4(4)	ABCDEFGHIJKLMNOPQRST
		CCCACA(12/9)		
<i>Rsa</i> I		GT↓AC		ABCDEFGHIJLMNOPQRST
	<i>Afa</i> I	GT↓AC		K
	<i>Csp</i> 6I	G↓TAC		DFN
<i>Rsr</i> II		CG↓GWCCG		BMN
	<i>Cpo</i> I	CG↓GWCCG		ADFK
	<i>Csp</i> I	CG↓GWCCG		EOR
<i>Sac</i> I		GAGCT↓C	5(?)	ACEFGHIJKLMNOPQRST
	<i>Ecl</i> 136II	GAG↓CTC		DFN

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>Eco</i> ICRI	GAG↓CTC		R
	<i>Psp</i> 124BI	GAGCT↓C		DI
	<i>Sst</i> I	GAGCT↓C		BC
<i>Sac</i> II		CCGC↓GG		ACEGHJKLNOPQRT
	<i>Cfr</i> 42I	CCGC↓GG		DF
	<i>Ksp</i> I	CCGC↓GG		MS
	<i>Sfr</i> 303I	CCGC↓GG		I
	<i>Sst</i> II	CCGC↓GG		B
	( <i>Mlu</i> 113I)	CC↓GCGG		
<i>Sal</i> I		G↓TCGAC	5(6)	ABCDEFGHIJKLMNQRST
<i>San</i> DI		GG↓GWCCC		E
<i>Sap</i> I		GCTCTTC(1/4)		N
( <i>Sau</i> I)	<i>Aoc</i> I	CC↓TNAGG		M
	<i>Axy</i> I	CC↓TNAGG		J
	<i>Bse</i> 21I	CC↓TNAGG		I
	<i>Bsu</i> 36I	CC↓TNAGG		ENR
	<i>Cvn</i> I	CC↓TNAGG		B
	<i>Eco</i> 81I	CC↓TNAGG		ADFKO
<i>Sca</i> I		AGT↓ACT		ABCDEFGHIJKLNOPQRS
	<i>Acc</i> 113I	AGT↓ACT		I
	<i>Eco</i> 255I	AGT↓ACT		F
<i>Scr</i> FI		CC↓NGG	?(5)	DJMNOS
	<i>Bss</i> KI	↓CCNGG		N
	<i>Msp</i> R9I	CC↓NGG		I
<i>Sdu</i> I		GDGCH↓C		F
	<i>Bmy</i> I	GDGCH↓C		M
	<i>Bsp</i> 1286I	GDGCH↓C		ADJKNR
( <i>Sec</i> I)	<i>Bsa</i> II	C↓CNNGG		N
	<i>Bse</i> DI	C↓CNNGG		DF
<i>Sex</i> AI		A↓CCWGGT		M
<i>Sfa</i> NI		GCATC(5/9)		IN
	( <i>Bsc</i> AI)	GCATC(4/6)		
( <i>Sfe</i> I)	<i>Bfm</i> I	C↓TRYAG		F
	<i>Bst</i> SFI	C↓TRYAG		DI
	<i>Sfc</i> I	C↓TRYAG		N
<i>Sfi</i> I		GGCCNNNN↓NGGCC	?(4)	ABCDEFGHIJKLMNQRST
<i>Sgf</i> I		GCGAT↓CGC		R
<i>Sgr</i> AI		CR↓CCGGYG		M
( <i>Sim</i> I)		GGGTC(-3/0)		
<i>Sma</i> I		CCC↓GGG	2(4)	ABCDEFGHIJKLMNQRST
	<i>Cfr</i> 9I	C↓CCGGG	2(4)	FO
	<i>Psp</i> AI	C↓CCGGG		E
	<i>Xma</i> I	C↓CCGGG	?(4)	ADINPR
	<i>Xma</i> CI	C↓CCGGG		M
<i>Sml</i> I		C↓TYRAG		N
( <i>Sna</i> I)	<i>Bss</i> NAI	GTA↓TAC		I
	<i>Bst</i> 1107I	GTA↓TAC		DFKM
	<i>Bst</i> Z17I	GTA↓TAC		N
<i>Sna</i> BI		TAC↓GTA		ACEKLMNR
	<i>Bst</i> SNI	TAC↓GTA		DI

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
	<i>Eco105I</i>	TAC↓GTA		FO
<i>SpeI</i>		A↓CTAGT		ABDEHIKLMNOPQRST
	<i>AcI</i> NI	A↓CTAGT		I
	<i>BcuI</i>	A↓CTAGT		F
<i>SphI</i>		GCATG↓C		ABCDEFGHIJKLMNPQRST
	<i>BbuI</i>	GCATG↓C		R
	<i>PaeI</i>	GCATG↓C		F
<i>SplI</i>		C↓GTACG		AK
	<i>Bst</i> WI	C↓GTACG		MNO
	<i>Pfl</i> 23II	C↓GTACG		DF
	<i>Psp</i> LI	C↓GTACG		I
	<i>SunI</i>	C↓GTACG		B
<i>SrfI</i>		GCCC↓GGGC		EO
<i>Sse</i> 8387I		CCTGCA↓GG		AK
	<i>SbfI</i>	CCTGCA↓GG		DIN
	<i>SdaI</i>	CCTGCA↓GG		F
( <i>Sse</i> 8647I)		AG↓GWCCT		
<i>SspI</i>		AAT↓ATT	?(6)	ABCDEFGHIJKLMNPQRS
<i>StuI</i>		AGG↓CCT		ABEJKLMNPQRST
	<i>AatI</i>	AGG↓CCT		O
	<i>Eco</i> 147I	AGG↓CCT		DF
	<i>Pme</i> 55I	AGG↓CCT		I
	<i>Sse</i> BI	AGG↓CCT		C
<i>StyI</i>		C↓CWWGG		BCEJMNRS
	<i>Bss</i> T1I	C↓CWWGG		I
	<i>Eco</i> 130I	C↓CWWGG		DF
	<i>Eco</i> T14I	C↓CWWGG		AK
	<i>ErhI</i>	C↓CWWGG		I
<i>SwaI</i>		ATTT↓AAAT		MS
	<i>SmiI</i>	ATTT↓AAAT		DFI
<i>TaqI</i>		T↓CGA	4(6)	ABCDEFGHIJLMNOPQRST
	<i>Tth</i> HB8I	T↓CGA	4(6)	AK
( <i>Taq</i> II) <sup>h</sup>		CACCCA(11/9)		
( <i>Taq</i> II) <sup>h</sup>		GACCGA(11/9)		
<i>TatI</i>		W↓GTACW		F
( <i>Tau</i> I)		GCSGC		
<i>TfiI</i>		G↓AWTC		N
<i>TseI</i>		G↓CWGC		N
<i>Tsp</i> 45I		↓GTSAC		N
( <i>Tsp</i> 4CI)	<i>Bst</i> 4CI	ACN↓GT		I
<i>Tsp</i> EI		↓AATT		LO
	<i>Sse</i> 9I	↓AATT	2(6)	DI
	<i>Tsp</i> 509I	↓AATT		N
<i>Tsp</i> RI		CASTGNN↓		N
<i>Tth</i> 111I		GACN↓NNGTC		ADIKNQR
	<i>Asp</i> I	GACN↓NNGTC		MS
	<i>Pfl</i> FI	GACN↓NNGTC		N
( <i>Tth</i> 111II)		CAARCA(11/9)		
( <i>Uba</i> DI)		GAACNNNNNTCC		
( <i>Uba</i> EI)		CACCTGC		

Table 1. continued

Prototype <sup>a</sup>	Isoschizomers <sup>a</sup>	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>	Commercial source <sup>d</sup>
<i>VspI</i>		AT↓TAAT	5(6)	BDEFIR
	<i>AseI</i>	AT↓TAAT		AJNO
	<i>AsnI</i>	AT↓TAAT		MS
	<i>PshBI</i>	AT↓TAAT		K
<i>XbaI</i>		T↓CTAGA	6(6)	ABCDEFGHIJKLMNQRST
<i>XcmI</i>		CCANNNN↓NNNTGG		AN
<i>XhoI</i>		C↓TCGAG	5(6)	ABCDEFGHIJKLMNQRST
	<i>PaeR7I</i>	C↓TCGAG		N
	( <i>SciI</i> )	CTC↓GAG		
	<i>Sfr274I</i>	C↓TCGAG		I
<i>XhoII</i>		R↓GATCY		EMRS
	<i>BstX2I</i>	R↓GATCY		DI
	<i>BstYI</i>	R↓GATCY		BN
	<i>MfiI</i>	R↓GATCY		AK
<i>XmaIII</i>		C↓GGCCG	4(5)	BE
	<i>BseX3I</i>	C↓GGCCG		I
	<i>BstZI</i>	C↓GGCCG	R	
	<i>EagI</i>	C↓GGCCG	?(5)	NP
	<i>EclXI</i>	C↓GGCCG		MS
	<i>Eco52I</i>	C↓GGCCG		ADFKOR
<i>XmnI</i>		GAANN↓NNTTC		DENR
	<i>Asp700I</i>	GAANN↓NNTTC		M
	<i>MroXI</i>	GAANN↓NNTTC		I

**Type I restriction enzymes**

Prototype enzyme	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>
<i>CfrAI</i>	GCANNNNNNNGTGG	
<i>EcoAI</i>	GAGNNNNNNGTCA	2(6),-3(6)
<i>EcoBI</i>	TGANNNNNNNTGCT	3(6),-4(6)
<i>EcoDI</i>	TTANNNNNNNGTCY	
<i>EcoDR2</i>	TCANNNNNNNGTCG	
<i>EcoDR3</i>	TCANNNNNNNATCG	
<i>EcoDXXI</i>	TCANNNNNNNRTTC	
<i>EcoEI</i>	GAGNNNNNNNATGC	
<i>EcoKI</i>	AACNNNNNNGTGC	2(6),-3(6)
<i>EcoR124I</i>	GAANNNNNNRTCG	3(6),-3(6)
<i>EcoR124II</i>	GAANNNNNNRTCG	3(6),-3(6)
<i>EcoRD2</i>	GAANNNNNNRTTC	
<i>EcoRD3</i>	GAANNNNNNRTTC	
<i>EcoprrI</i>	CCANNNNNNRTGC	
<i>StyLTIII</i>	GAGNNNNNNRTAYG	2(6),-4(6)
<i>StySJ</i>	GAGNNNNNNGTRC	
<i>StySKI</i>	CGATNNNNNNGTTA	
<i>StySPI</i>	AACNNNNNNGTRC	2(6),-3(6)
<i>StySQ</i>	AACNNNNNNRTAYG	

**Type III restriction enzymes**

Prototype enzyme	Recognition sequence <sup>b</sup>	Me site <sup>c</sup>
<i>EcoPI</i>	AGACC	3(6)
<i>EcoP15I</i>	CAGCAG(25/27)	5(6)
<i>HinfIII</i>	CGAAT	
<i>StyLTI</i>	CAGAG	4(6)

<sup>a</sup>When an enzyme name appears in parentheses it indicates that it is not itself commercially available. \*Signifies that *DpnI* requires the presence of 6-methyladenosine within the recognition sequence GmATC.

<sup>b</sup>Recognition sequences use the standard abbreviations [*Eur. J. Biochem.* (1985) **150**, 1–5] to represent ambiguity.

R = A or G            B = not A (C or G or T)  
 Y = C or T            D = not C (A or G or T)  
 M = A or C            H = not G (A or C or T)  
 K = G or T            V = not T (A or C or G)  
 S = G or C            N = A or C or G or T  
 W = A or T

Recognition sequences are written from 5' to 3', only one strand being given, and the point of cleavage is indicated by an arrow (↓). When no arrow appears, the precise cleavage site has not been determined. For enzymes such as *BciVI*, which cleaves away from its recognition sequence, the sites of cleavage are indicated in parentheses. Thus GTATCC (6/5) indicates cleavage as shown below:

5'-GTATCCNNNNNN↓-3'  
 3'-CATAGNNNNNN↓-5'

In all cases the recognition sequences are oriented so that the cleavage sites lie on their 3'-side.

<sup>c</sup>The site of methylation by the cognate methylase when known is indicated as follows. The first number shows the base within the recognition sequence that is modified. A negative number indicates the complementary strand, numbered from the 5' base of that strand. The number in parentheses indicates the specific methylation involved. (6) = N6-methyladenosine; (5) = 5-methylcytosine; (4) = N4-methylcytosine. Thus for *Aflw26I* the designation 3(5),-4(6) signifies the following pattern of methylation:

1        2        3        4        5  
 5'    G        T        5mC     T        C    3'  
 3'    C        6mA     G        A        G    5'  
      -5       -4       -3       -2       -1

<sup>d</sup>Commercial sources of restriction enzymes are abbreviated as follows: A, Amersham Life Sciences-USB (4/97); B, Life Technologies Inc, Gibco-BRL (1/97); C, Minotech Molecular Biology Products (10/97); D, Angewandte Gentechnologie Systeme (10/97); E, Stratagene (4/97); F, Fermentas MBI (9/97); G, Appligene Oncor (10/97); H, American Allied Biochemical, Inc. (10/97); I, SibEnzyme Ltd. (4/97); J, Nippon Gene Co., Ltd. (10/97); K, Takara Shuzo Co. Ltd. (4/97); L, NBL Gene Sciences Limited (5/97); M, Boehringer-Mannheim (10/97); N, New England BioLabs (10/97); O, Toyobo Biochemicals (7/97); P, Pharmacia Biotech Inc. (8/97); Q, CHIMERx (10/97); R, Promega Corporation (7/97); S, Sigma (10/97); T, Advanced Biotechnologies Ltd. (1/97). The date in parentheses indicates the most recent update.

<sup>e</sup>Enzymes that cut on both sides of their recognition sequence, such as *BaeI*, *BcgI*, *Bsp24I*, *CjeI* and *CjePI*, have four cleavage sites each instead of two. For example, *Bsp24I* cleaves as indicated:

5'                            ↓NNNNNNNNGACNNNNNNNTGGNNNNNNNNNNNN↓            3'  
 3'                            ↓NNNNNNNNNNNNNCTGNNNNNNNACCNNNNNNNN↓            5'

This is described in Table 1 and some REBASE reports as *Bsp24I* (8/13)GACNNNNNNNTGG(12/7). Thus the recognition site is excised in a fragment, 32 bp long, with 5-base 3'-extensions at each end.

<sup>f</sup>*EcoRII* isoschizomers fall into two classes based upon their sensitivity to methylation. *EcoRII* will not cleave when the second cytosine in the recognition sequence is methylated to 5-methylcytosine whereas *MvaI* will cleave such a sequence. Isoschizomers of *EcoRII* that are like *MvaI* are indicated by +.

<sup>g</sup>*MboI* isoschizomers fall into two classes based upon their sensitivity to methylation. *MboI* will not cleave when the recognition sequence contains 6-methyladenosine whereas *Sau3AI* will not cleave when its recognition sequence contains 5-methylcytosine. Isoschizomers of *MboI* that are like *Sau3AI* are indicated by +.

<sup>h</sup>*TaqII* differs from other restriction enzymes in recognizing two distinct sequences: GACCG and CACCCA.

**Table 2.** Homing endonucleases

Enzyme	Sequence around one cleavage site <sup>a</sup>	Commercial source <sup>b</sup>
I-CeuI	CGTAACTATAACGGTC_CTAA↓GGTAGCGAA	N
I-DmoI	ATGCCTTGCCGG_GTAA↓GTTCCGGCGCGCAT	L
I-PpoI	TAACTATGACTCTC_TTAA↓GGTAGCCAAAT	NR
I-SceI	AGTTACGCTAGGG_ATAA↓CAGGGTAATATAG	M
PI-PspI	AAAATCCTGGCAAACAGCTA_TTAT↓GGGTAT	N
PI-SceI	ATCTATGTCGG_GTGC↓GGAGAAAGAGGTAAT	N
PI-TliI	GTTCTTTATGCGG_ACAC↓TGACGGCTTTTA	N

<sup>a</sup>The recognition sequence is loosely defined as the sequence that normally flanks the intron or intein encoding the endonuclease. ↓ indicates cleavage on the strand shown; \_ indicates cleavage on the complementary strand.

<sup>b</sup>For commercial sources see the legend to Table 1.